Error Analysis for SEE Cross Sections

Ray Ladbury
NASA/GSFC
Radiation Effects and Analysis Group

What!? No Tales from the Cave?

- Look at one of the Lessons Learned: Assigning Error Bars

I. What kinds of errors are important for SEE testing?
   A. Random Errors
   B. Systematic Errors

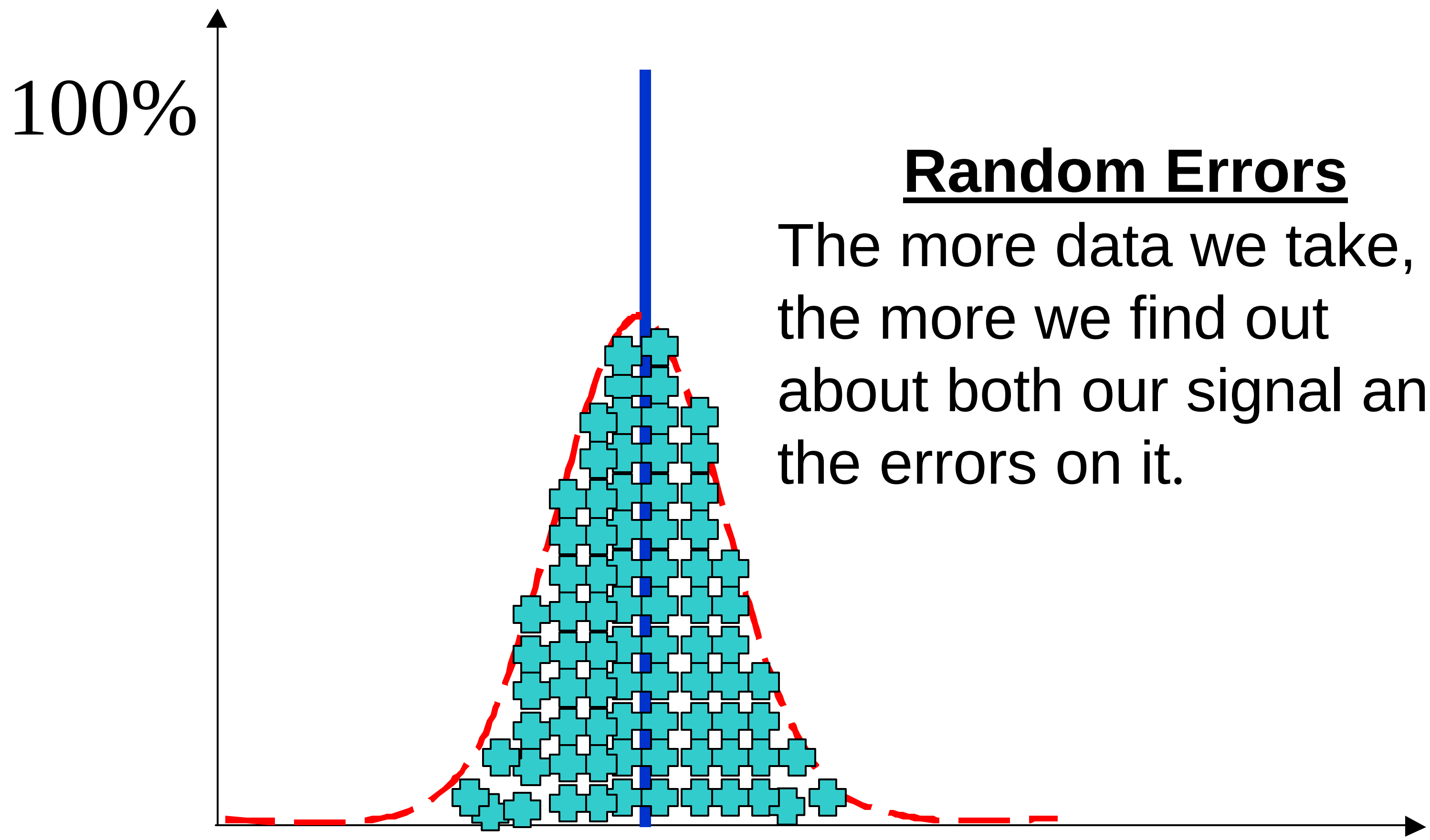
II. Estimating Random Errors
   A. Poisson fluctuations of SEE counts
   B. Part-to-Part variations (Binomial statistics...or not?)
   C. Others?

III. Distribution-Independent Error Analysis
   A. Bootstrapping
   B. Examples

IV. Systematic Errors—a work in progress
   A. Contamination of Datasets
   B. Others

V. Conclusions

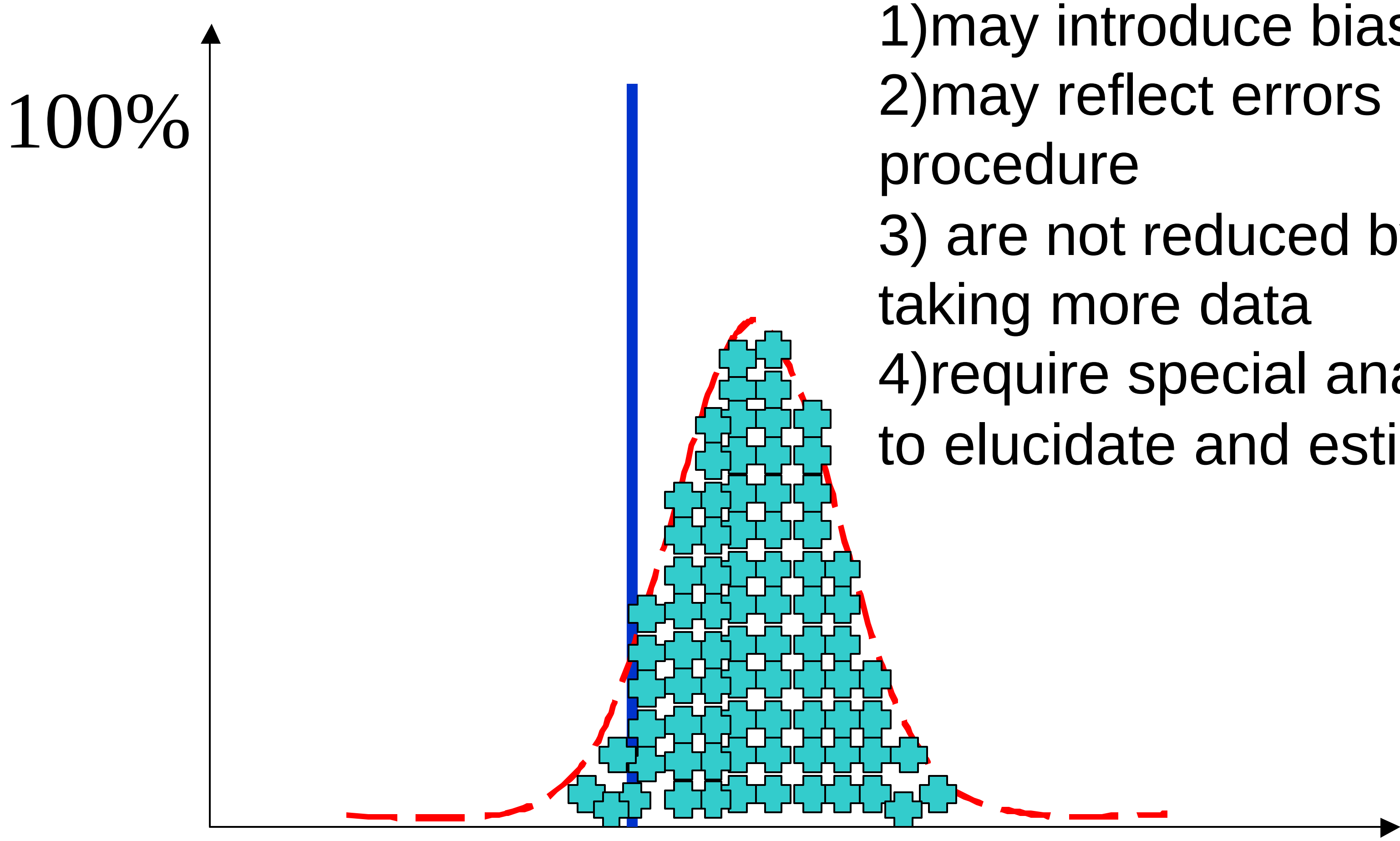
Random Errors



Random Errors

The more data we take, the more we find out about both our signal and the errors on it.

Systematic Errors



Systematic Errors

1)may introduce biases

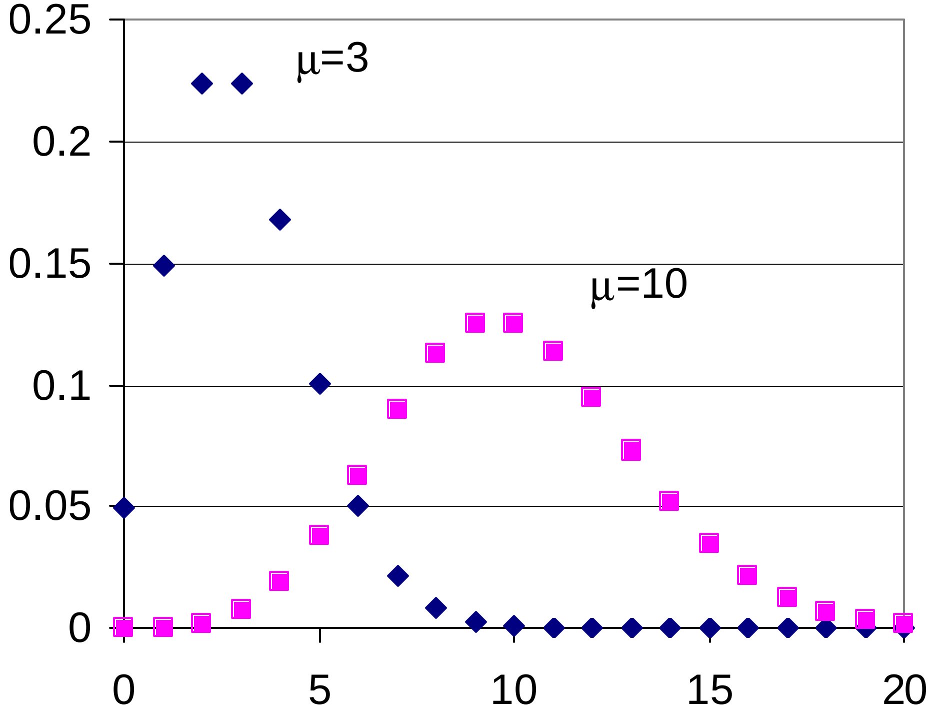
2)may reflect errors in procedure

3) are not reduced by taking more data

4)require special analysis to elucidate and estimate

Random Errors: Poisson Fluctuations

- Poisson distribution—probability of *n* counts when we expect μ
  - $P(n, \mu)= \mu^n e^{\mu}/n!$
  - asymmetric, esp. for μ small
  - standard deviation, sd= $\mu^{1/2}$



- Observation of *n* SEE counts may represent a fluctuation from the real mean μ
  - real cross section $\sigma= \mu$/fluence
- What can we say about μ if we observe *n* events?
  - Look at what μ could be if our observation of n just barely has probability 1-CL

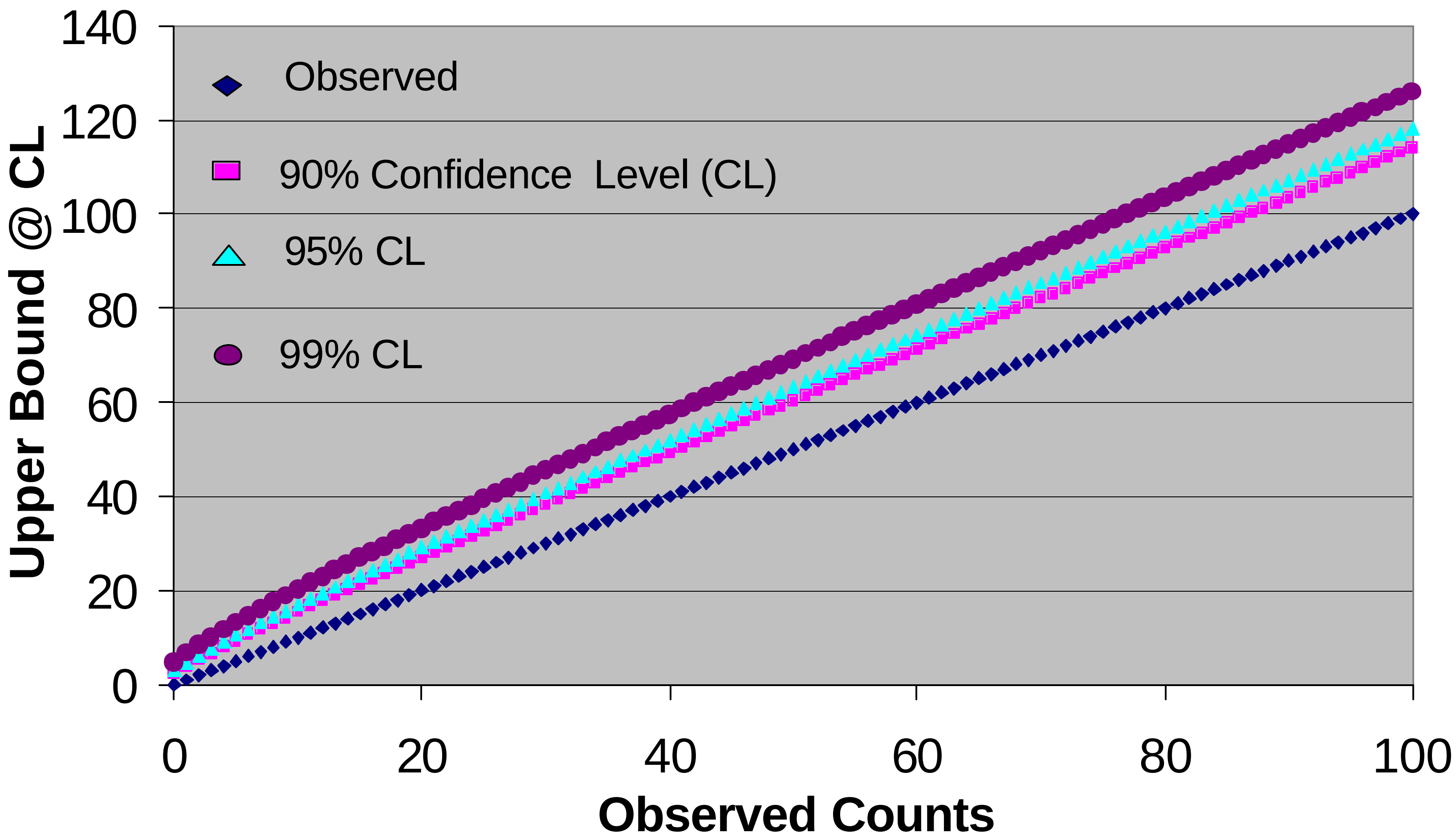
Upper Bound for mean @ CL

Confidence Level	OBSERVED			
	0	1	2	3
90%	2.305	3.89	5.32	6.68
95%	2.996	4.74	6.3	7.75
99%	4.605	6.64	8.41	10.04

Upper Bounds for μ Given *n*

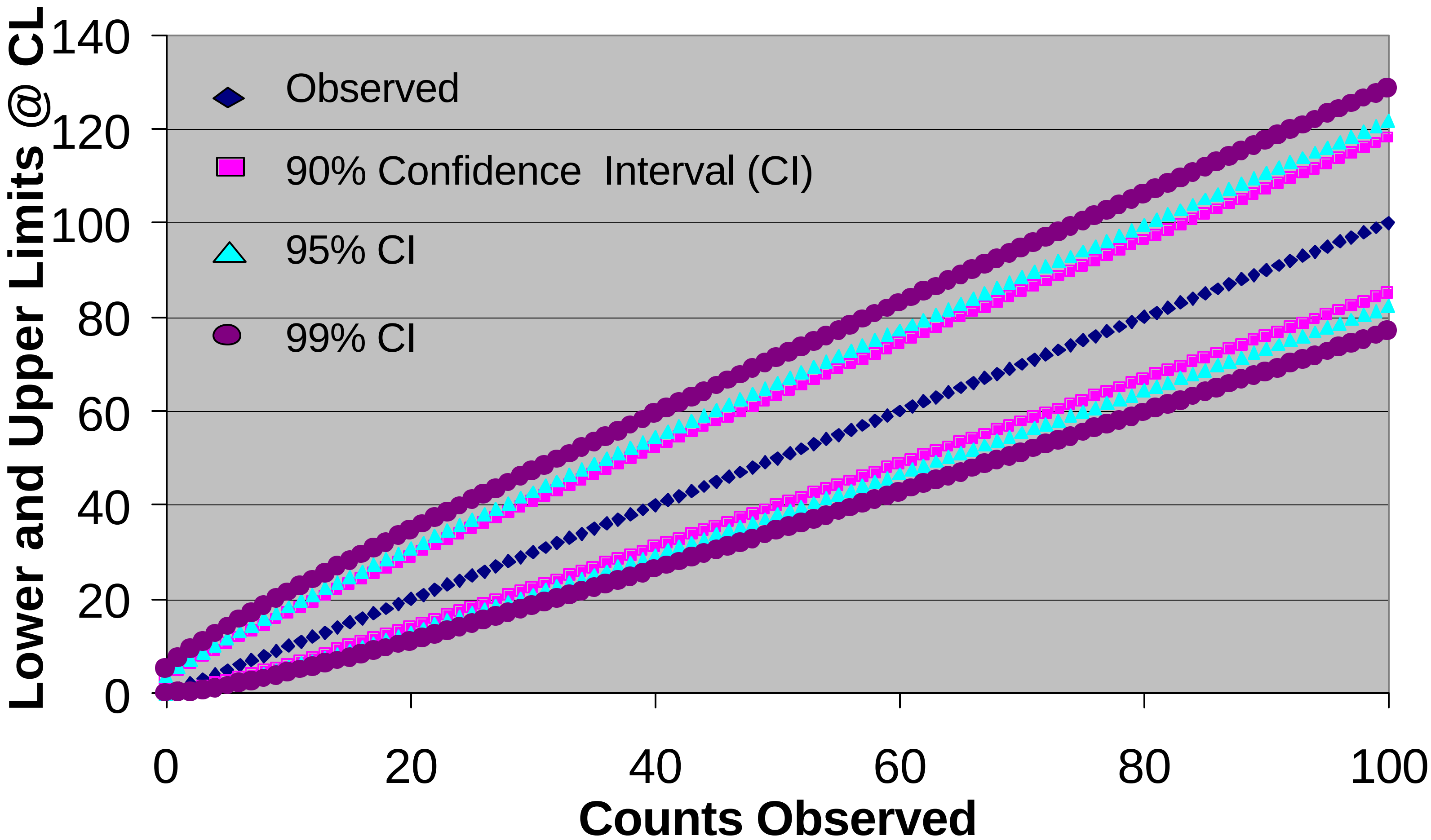
Upper bound useful for computing bounding rates using FOM approach

$$R_{FOM} = \frac{C\sigma_{lim}}{LET_0^2}$$



Confidence Interval for μ Given *n*

Confidence intervals (upper and lower bounds for μ consistent with a given confidence CL) can be used to define upper and lower error bars for SEE cross section measurements.



Other Random Errors

Part-to-part variability

- Have not been major concerns in SEE testing
- But...
  - Commercial parts sometimes show lot-to-lot variation
    - How do we know our test sample is representative?
- Dealing with variation
  - Binomial Statistics
    - Independent of distribution but requires large samples
  - Assume a distribution form
    - Less general, but requires smaller samples

- Others?
  - Noise (e.g. on an A to D)
  - Measurement errors
    - Normally distributed?
  - Beam fluctuations?
    - probably random; if systematic, they would be noticed facility to facility
  - And so on
- So we could have
  - Poisson errors on event counts
  - Sampling errors part-to-part
  - Various other errors
- This is getting complicated!
- Can we use a distribution-independent method

Bootstrapping

- Bootstrapping constructs a distribution from the samples
  - Suppose we have a sample from a parent distribution:

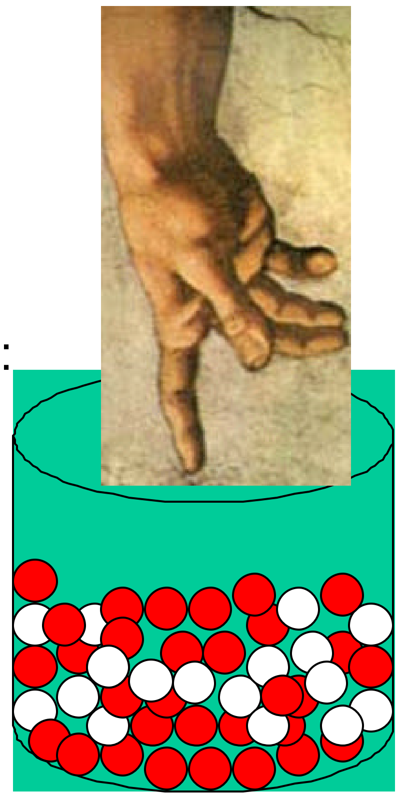
  $\{n_1, n_2, n_3, \ldots, n_{m-1}, n_m\}$

  - Construct a large number of m-element samples by drawing with replacement:

  $\{n_5, n_1, n_5, \ldots, n_2, n_3\}$

  $\{n_7, n_2, , n_4, \ldots, n_{m-6}, n_{m-9}\}$ and so on

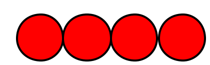
  - Bootstrapped sample statistics reasonably reproduce those of the parent distribution
    - no assumptions about distribution



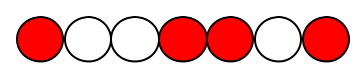
- Caveats:
- As Tom Lehrer's friend Hen3ry said:

"Life is like a sewer; you get out of it what you put into it."

- Bootstrapping analyses are only as representative as the samples on which they are based.



not representative



more representative

Bootstrapping Errors for SEU

Example: Errors on σ_{SEU} for an SDRAM

1^{st} read:
n_1 errors after f_1 ions

after run of m reads:
$\{(n_1,f_1), (n_2,f_2)\ldots (n_m,f_m)\}$

$\sigma = (n_1+n_2+\ldots+n)/(f_1+f_2+ f_m)$

make pseudoruns from
$\{(n_1,f_1), (n_2,f_2)\ldots (n_m,f_m)\}$,
each with m "reads"

$\{(n_3,f_3), (n_m,f_m)\ldots (n_m,f_m)\}$
$\{(n_2,f_2), (n_2,f_2)\ldots (n_5,f_5)\}$
.
.
.
$\{(n_6,f_6), (n_4,f_4)\ldots (n_9,f_9)\}$

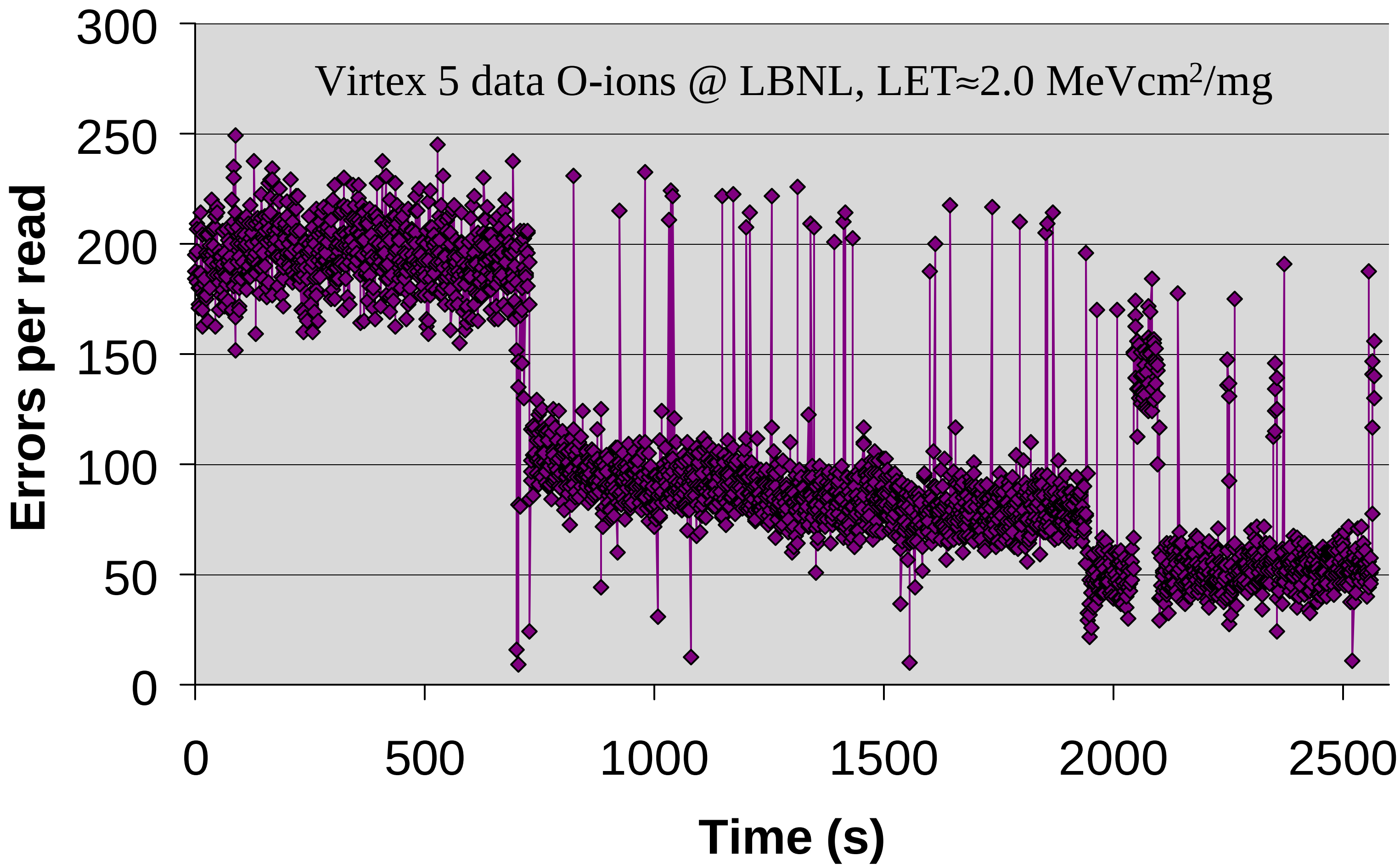
calculate σ^*s=
$$\frac{(n_1^*+n_2^*+\ldots+n_m^*)}{(f_1^*+f_2^*+ f_m^*)}$$
for each pseudorun

Rank σ^*s smallest to largest:
If we have 10000 σ^*s,
σ^*_{9000} is the 90% CL upper bound on σ;
σ^*_{500} and σ^*_{9500} are the limits of the 90% CI for σ

Note: No assumptions made about distribution of σs—the data determine the distribution.
Whatever we include gets modeled:
Poisson fluctuations
Part-to-part variations
etc.

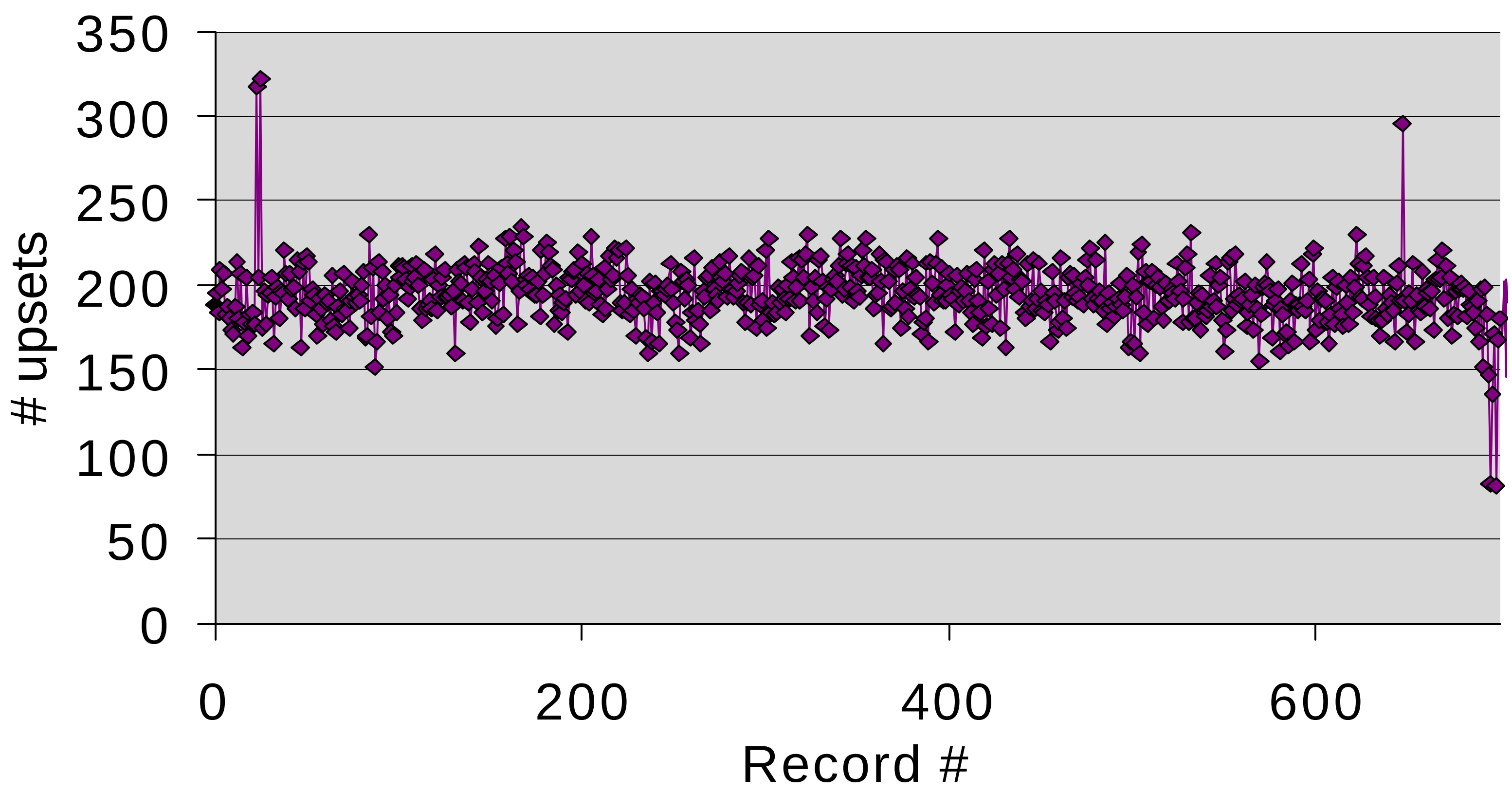
Yeah, But Will it Work on Real Data?

- Clearly need clean data to look at each SEE (but need that anyway)
- Need to know the real fluence and time for each read
  - Could we use the moving average + record # to estimate flux vs. time?



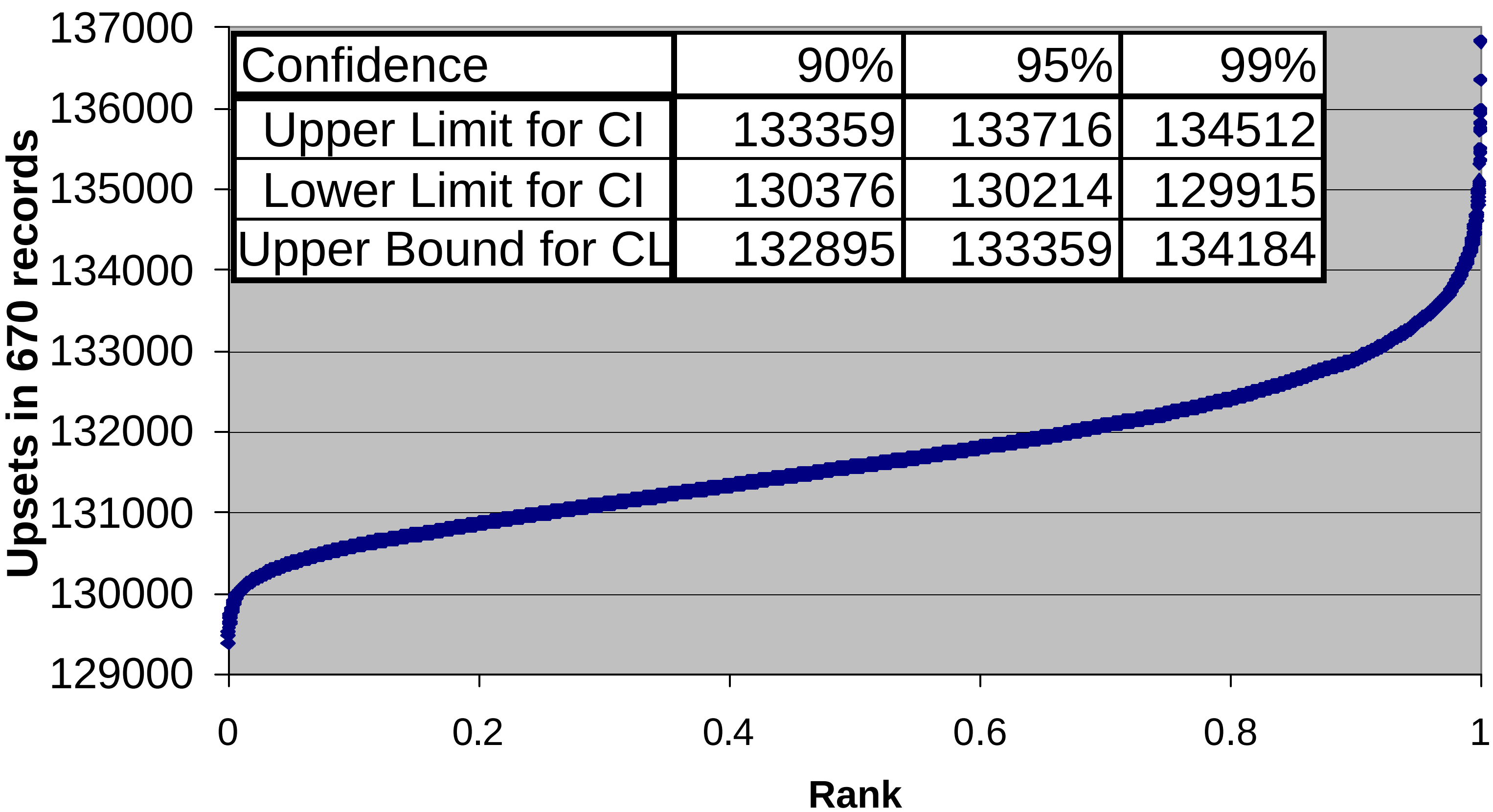
Cleaner Data

- The nice thing about big errors is they stand out
  - The cleaned dataset keeps most of the SEUs and tosses most of the SEFIs
  - We look only at the first ~680 records where we're pretty sure things are operating as expected.



Bootstrap for "Clean" Data

- Look at 1st 670 records where performance is consistent
  - (130882 SEUs)
- Generate 10000 pseudoruns of 670 records each
  - Rank event totals for pseudoruns smallest to largest



Confidence	90%	95%	99%
Upper Limit for CI	133359	133716	134512
Lower Limit for CI	130376	130214	129915
Upper Bound for CL	132895	133359	134184

Systematic Errors

- Contamination of SEU data with SEFIs is a systematic error
  - leads to overestimate of the cross section
  - may become more significant at high LET
  - taking more data does not reduce the errors
- Systematic errors need to be investigated and estimated
  - special experiments or analyses are needed
- Other possible systematic errors for SEE cross sections
  - miscalibration of fluence, dead time in experiment, burst errors
- Systematic errors may also have some distribution
- Combining random and systematic errors (use same CL for both)
  - Independent of each other, so best estimate is RMS
  - Bounding estimate is the sum of the absolute values
  - Important: Systematic errors usually not symmetric—combine by sign

$$R \pm_{b}^{a} \pm_{d}^{c} \qquad R \pm_{RMS(b,d)}^{RMS(a,c)} \qquad R \pm_{b+d}^{a+c}$$

Conclusions

- Error analysis for SEE is complicated
  - May have multiple sources of both random and systematic errors
  - Systematic errors especially may not be well understood
- Random errors are assumed to be Poisson
  - If there may be other sources of random errors—bootstrapping provides a distribution-independent approach for error analysis
    - Need to know ion fluence for each readout
- Systematic errors remain a challenge
  - Contamination, deadtime, burst errors may all become more important as parts become more complicated.
- If we can model random and systematic errors, we can bound SEE rates for a given CL and consistent with experimental limitations
  - SEE data is always limited
  - When testing complicated parts, it may be even more so.